

The Impact of Intra-Die Device Parameter Variations on Path Delays and on the Design for Yield of Low Voltage Digital Circuits

Martin Eisele, *Student Member, IEEE*, Jörg Berthold, Doris Schmitt-Landsiedel, and Reinhard Mahnkopf

Abstract—The yield of low voltage digital circuits is found to be sensitive to local gate delay variations due to uncorrelated intra-die parameter deviations. Caused by statistical deviations of the doping concentration they lead to more pronounced delay variations for minimum transistor sizes. Their influence on path delays in digital circuits is verified using a carry select adder test circuit fabricated in 0.5 and 0.35 μm complementary metal-oxide-semiconductor (CMOS) technologies with two different threshold voltages. The increase of the path delay variations for smaller device dimensions and reduced supply voltages as well as the dependence on the path length is shown. It is found that circuits with a large number of critical paths and with a low logic depth are most sensitive to uncorrelated gate delay variations. Scenarios for future technologies show the increased impact of uncorrelated delay variations on digital design. A reduction of the maximal clock frequency of 10% is found for e.g. highly pipelined systems realized in a 0.18- μm CMOS technology.

Index Terms—Gate delay variations, low-voltage digital design, path delay variations, yield, parameter variations, pipelined circuits, scaling, SRAM, V_{th} variations.

I. INTRODUCTION

FOR an increasing range of applications, low power dissipation is an integrated circuit (IC) feature as important as area and speed. In the case of portable systems, low power circuits allow longer periods of operation. For IC's embedded in high performance systems, a low power dissipation reduces the expenses for chip cooling or increases the reliability due to the lower chip temperature.

The main contribution to the power dissipation is the dynamic power P_{dyn} for charging the capacitive loads. According to

$$P_{dyn} \propto C_{load} \cdot V_{DD}^2 \quad (1)$$

in the case of static CMOS circuits, where V_{DD} is the supply voltage and C_{load} is the load capacitance (including the capacitances for the wiring, the junctions, and the transistor gates), the power dissipation can be significantly reduced by scaling the supply voltage. Additionally, power can be saved

by reducing the capacitive load C_{load} with, e.g., using small device sizes.

The reduction of the supply voltage faces two disadvantages; first, the gate delays are increased. To first order, the gate delay d_{gate} of an inverter can be described by

$$d_{gate} \propto \frac{C_{load} V_{DD}}{\mu C_{ox} (W/L) (V_{DD} - V_{th})^\alpha} \quad (2)$$

where μ is the mobility, V_{th} is the threshold voltage, C_{ox} is the gate capacitance per unit area, W and L are the transistor dimensions, and α has a value between two and one, depending on the short channel effect [1]. Second, a reduced supply voltage leads to an increased sensitivity $S_f^{V_{th}}$ of gate delays to parameter variations influencing V_{th} according to

$$S_f^{V_{th}} = \frac{V_{th}}{d_{gate}} \cdot \frac{\partial d_{gate}}{\partial V_{th}} = \frac{\alpha \cdot V_{th}}{(V_{DD} - V_{th})} \quad (3)$$

Furthermore, transistors with small gate sizes increase the effect of geometry dependent parameter variations.

For conventional supply voltages, i.e., $V_{DD} > 3-4 \times V_{th}$ only *interdie* parameter variations are to be considered in digital designs. These variations act *globally* on the entire chip or on functional blocks, so that each device on one chip or in one block shows the same deviation. Inter-chip or inter-block variations can be caused by systematic effects like process gradients over the wafer [2] with typical distances in the range of functional block sizes or above. Variations of the gate oxide thickness can, e.g., be regarded as global variations. Sets of worst-case and best-case parameters are used during design verifications to cover the influence of these global variations. Effects where each device on the chip is affected differently are called *intra-die* or *local* variations. They are overlaid to the global variations and are caused by e.g. proximity effects or statistical variations of doping concentrations. For current technologies using standard device sizes, they are small. Therefore, for the design of digital circuits operating at conventional supply voltages they are regarded to be negligible.

The impact of global parameter variation on the performance and power dissipation is investigated in, e.g., [3] and [2], [4], respectively. Since small dimension devices operating at low supply voltages show an increased sensitivity to parameter variations the impact of local variations may become significant. This has to be investigated for issues like skew sensitive signals, testing methods, design for yield,

Manuscript received September 8, 1996; revised May 15, 1997.

M. Eisele is with the Institute of Technical Electronics, Technical University of Munich, Germany and Siemens Corporate R&D, Munich 81730 Germany.

J. Berthold is with the Siemens Corporate R&D, Munich 81730 Germany.

D. Schmitt-Landsiedel is with the Institute of Technical Electronics, Technical University of Munich, Munich 80290 Germany.

R. Mahnkopf is with the Siemens Semiconductor Division, Munich 81730 Germany.

Publisher Item Identifier S 1063-8210(97)06422-6.

and performance estimation. Therefore, for the design of low voltage digital circuits the effect of intra-die local parameter variations has to be quantified.

In this paper, the influence of local V_{th} -variations on path delays for different supply voltages, path lengths, transistor sizes, and technologies is investigated. Measurements are performed on a carry select adder test circuit realized in a 0.5 and 0.35 μm CMOS technology to verify this influence. To demonstrate the impact on the design of low voltage digital circuits, the influence on yield and on performance estimation of pipelined systems is studied. To account for future CMOS technologies the influence of local variations on the circuit performance is calculated using technology parameters according to the SIA (Semiconductor Industry Association) road map [5].

The paper is organized as follows. In Section II, data of intra-die V_{th} -variations are presented and their impact on gate delays and path delays is given. The test circuit and the measurement results of path delay variations are shown in Section III. In Section IV, the increase of gate delay variations in future technology parameters is quantified. The impact on the design is shown for yield evaluation and performance estimation of pipelined architectures and SRAM's in Section 5. The results are summarized in the conclusion. The detailed derivations of the gate and path delay variances are given in the Appendix.

II. V_{th} -VARIATIONS AND THEIR IMPACT ON GATE AND PATH DELAYS

According to (2), the main parameter leading to an increase of the sensitivity S is the threshold voltage V_{th} . Therefore, investigations of V_{th} -variations of small digital transistors are of high importance in the design of low voltage digital circuits [6], [7]. Previous measurements in [6] show uncorrelated intra-die V_{th} -variations with the $1/\sqrt{WL_{eff}}$ -dependence (Fig. 1). Even transistors with quite different W/L_{eff} ratios like 0.7/4.4, 1.8/1.7, and 7.5/0.4 but with the same gate area $1/\sqrt{WL_{eff}} \approx 0.57 \mu\text{m}^{-1}$ are showing similar standard deviations. Therefore, the variations are dominated by an area dependent parameter. Variations of the oxide thickness t_{ox} can be excluded due to the effect, that variations of V_{th} results in parallel shifts of the input characteristics. Moreover, to detect proximity effects different surroundings in polysilicon coverage and field oxide areas were investigated but are shown to be negligible for a 0.5 μm technology. This means, that the V_{th} -variations are dominated by the statistical deviations of the doping concentration in the channel area due to the implantation process [6]. They can be approximated with

$$\sigma_{V_{th}}^2 = \frac{Q_B^2}{4L_{eff}WW_dN} + \frac{q^2}{C_{OX}^2} \cdot \frac{D_i}{WL_{eff}} \quad (4)$$

where Q_B is the depletion charge density, W_d is the depletion layer width, N is the substrate doping, D_i is the channel doping implant dose, q the elementary charge, W is the transistor channel width, and L_{eff} is the effective channel length.

The influence of $\sigma_{V_{th}}$ on the delay time of single inverters with different transistor dimensions and supply voltages was

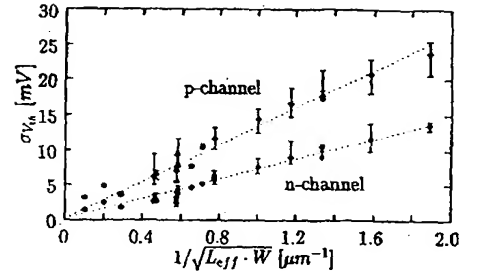


Fig. 1. Standard deviation of V_{th} versus $(WL_{eff})^{-1/2}$. The error bars are obtained from experimental results. The dotted lines represent least square fits to demonstrate the linear $(WL_{eff})^{-1/2}$ dependence.

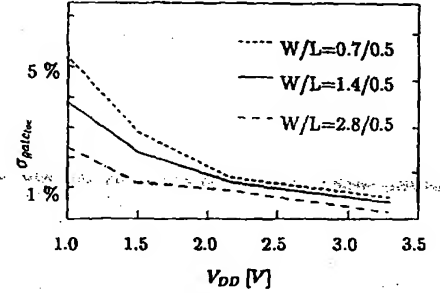


Fig. 2. Average delay variations of the low-high transitions and the high-low transitions ($W_p = 2W_n$) of single inverters with three different widths normalized to the nominal delay obtained from SPICE simulations using measured $\sigma_{V_{th}}$ values.

investigated by means of SPICE simulations using $V_{th} \pm 3\sigma_{V_{th}}$ values. The variation of the delay, normalized to the nominal delay is shown in Fig. 2 demonstrating the increase of delay variations at low supply voltages and smaller dimensions. The delay of an inverter realized with small transistors shows a relative delay variations due to local effects with $\sigma_{gate_{loc}}^2 = \sigma_{gate_{occ}}^2 / \bar{d}_{gate}^2 = 5\%$ or of $\sigma_{gate_{occ}}^2 = 3\%$ for $V_{DD} = 1.1 \text{ V}$ ($\approx 2.1 \times V_{th}$) and $V_{DD} = 1.5 \text{ V}$ ($\approx 2.9 \times V_{th}$), respectively.

The distance dependent correlation of the the gate delay on the chip is given by the autocorrelation coefficient $\rho_a(\delta)$ [8] with

$$\rho_a(\delta) = \frac{\sum (d_{gate}(x) - \bar{d}_{gate})(d_{gate}(x + \delta) - \bar{d}_{gate})}{\sum (d_{gate}(x) - \bar{d}_{gate})^2} \quad (5)$$

where $d_{gate}(x)$ is the delay of the x th gate in a path comprising n gates. As the variations are caused due to statistical deviations of the doping concentrations, no area or distance δ dependent correlation of the gate delays d_{gate} exist. The autocorrelation coefficient $\rho_a(\delta)$ can be set to $\rho_a(\delta) \approx 0$. This has to be compared with $\rho_a(\delta) \approx 1$ when only global effects are taken into account. With the variance of a single gate σ_{gate}^2 the variance of a path comprising n gates is given as [see Appendix, (17)]

$$\sigma_{path}^2 = \sum_{i=1}^n \sigma_{gate_i}^2 \quad (6)$$

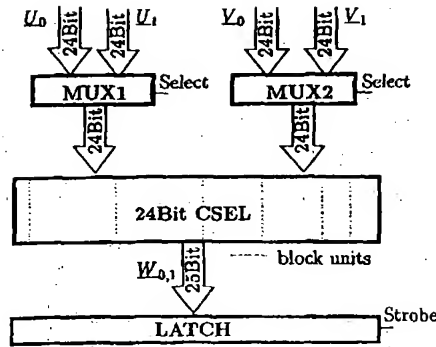


Fig. 3. Test structure of 24 bit carry select adder (2-2-4-4-6-6 block structure) for path delay measurements.

The relative delay variation of a path is reduced for longer paths when, e.g., using identical gates according to

$$\frac{\sigma_{\text{path}}}{d_{\text{path}}} \propto \frac{1}{\sqrt{n}} \quad (7)$$

To summarize the result of this section: First, local V_{th} variations are clearly dominated by the doping variations. For functional blocks with size in the order of the size of our test structure ($640 \times 640 \mu\text{m}$), global variations as, e.g., variations of oxide thickness, gate length, and others are of minor importance. Second, due to the local V_{th} variations an increase of the relative path delay variations is expected for shorter paths and for gates using small devices.

III. TEST CIRCUIT AND EXPERIMENTAL RESULTS

To verify the influence of the local V_{th} variations on logic circuits, a test circuit was designed, comprising a 24 bit carry select adder typically used in data path structures. The circuit is shown in Fig. 3 and is operated as follows: switching the signal *select* controlling the two multiplexer MUX1 and MUX2, new 24 bit patterns are applied to the adder inputs. The *strobe* signal allows the storage of the 25 bit result in LATCH. The skew between *strobe* and the *select* signal facilitate the precise determination of individual path delays.

The test circuit was fabricated with two different technologies. Firstly, with an $0.5 \mu\text{m}$ process with two threshold voltages: for wafers of type I, $V_{th} = 1/2(V_{thn} + V_{thp})$ equals 520 mV and for wafers of type II, V_{th} equals 290 mV. Secondly, wafers were processed with an $0.35 \mu\text{m}$ technology (type III) where V_{th} equals 400 mV. On each chip, the adder is realized using different transistor dimensions: for wafers of type I and II one with $W/L = 0.7 \mu\text{m}/0.5 \mu\text{m}$ and one with $W/L = 1.4 \mu\text{m}/0.5 \mu\text{m}$ for the smallest transistors used. For wafers of type III the ratios equal to $W/L = 0.5 \mu\text{m}/0.35 \mu\text{m}$ and to $W/L = 1.0 \mu\text{m}/0.35 \mu\text{m}$. For the *p*-channel devices a channel width of $W_p = 2 \times W_n$ is taken.

The design allows the concurrent measurement of identical, non overlapping, adjacent paths on the same chip. For the measurement of intra-die path delay deviations two different path lengths are chosen (Fig. 4): four paths having a logic depth of eight gates and 12 paths comprising four gates. Significant statistical data are obtained by the compilation of

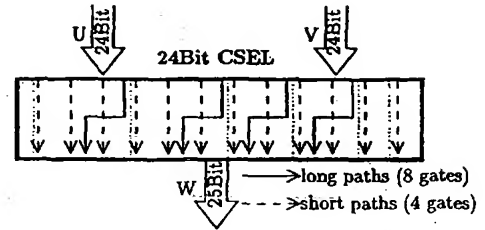


Fig. 4. Concurrent measurement of identical, non overlapping paths is possible, e.g., four paths with logic depth 8 or 12 paths with logic depth 4.

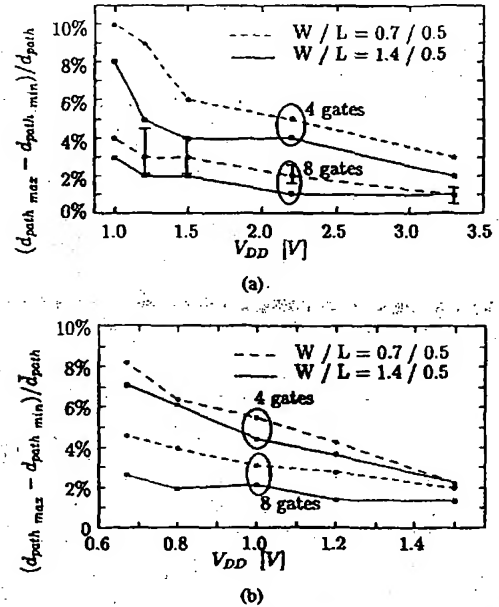


Fig. 5. Path delay variations normalized to the nominal delay for two different path lengths and transistor dimensions, showing the increased variations for lower supply voltages, smaller dimensions and reduced logic depth. The circuits in (a) of type I) are processed for $V_{th} = 520 \text{ mV}$ whereas in (b) for $V_{th} = 290 \text{ mV}$. Representative for one curve, the error bars are shown.

all path delay deviations normalized to the local mean value \bar{d}_{path} measured on several chips and wafers. Long distance effects are excluded.

The results are shown in Fig. 5 I) and II) for wafers of type I and II, respectively. For both wafer types, an increase of the relative path delay variations for smaller transistor dimensions and reduced supply voltages is found. For the same V_{DD} , the low V_{th} circuits (type II) show a reduced delay variation compared to type I. This is due to the higher value of $(V_{DD} - V_{th})$. Regarding the same V_{DD}/V_{th} ratios of both wafer types, nearly the same normalized delay variations are obtained (see Fig. 6). This can be explained by (4), where the sensitivity $S_{d_{\text{gate}}}^{V_{th}}$ only depends on the ratio of V_{DD}/V_{th} . The variations for both wafers reach a value of up to 9% for $(d_{\text{max}} - d_{\text{min}})/d_{\text{nom}}$ for a path with 4 gates operating at a supply voltage $V_{DD} = 2.1 \times V_{th}$ (Fig. 6).

The delay variations versus V_{DD}/V_{th} for wafers of type III are shown in Fig. 7. The same trend as with a $0.5 \mu\text{m}$ technology can be observed, but a higher value of the relative

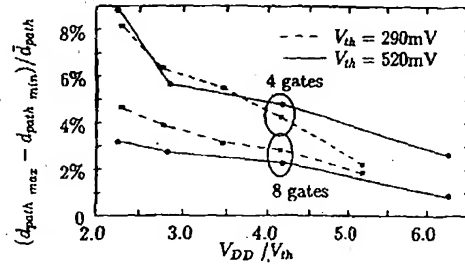


Fig. 6. The relative path delay variations versus V_{DD}/V_{th} show the similar values for structures realized with different threshold voltages.

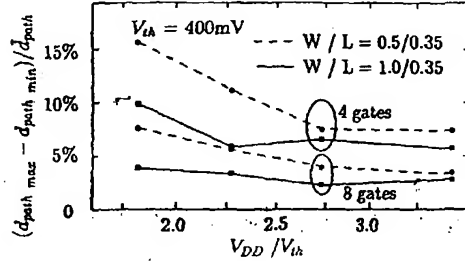


Fig. 7. The relative path delay variations versus V_{DD}/V_{th} for 0.35 μm technology. The variations are increased compared to a 0.5 μm technology due to increased V_{th} variations.

TABLE I
DEVICE TECHNOLOGY PARAMETERS, WHERE t_{OX} IS THE GATE OXIDE THICKNESS AND s IS THE SCALING FACTOR

Parameter	Technology			
	0.5 μm	0.35 μm	0.25 μm	0.18 μm
L_{eff} [μm]	0.4	0.28	0.20	0.14
t_{OX} [nm]	95	80	60	45
s	1	0.7	0.5	0.36

variations at the same V_{DD}/V_{th} ratio has to be regarded. This is due to increased V_{th} -variations for smaller devices and higher doping concentrations and will be explained in the next section.

The experimental results verify the assumptions of Section II. The relative path delay variations increase for smaller devices and reduced supply voltages and depend on the logic depth according to (7) confirming the domination of the local V_{th} -variations.

IV. IMPACT OF SCALING ON DELAY VARIATIONS

The observation that the delays of adjacent paths can deviate significantly has to be regarded in the design of digital low voltage circuits for actual and, moreover, for future CMOS technologies where the variations tend to increase. To account for future CMOS technologies the influence of local variations on the circuit performance is calculated using technology parameters according to the SIA road map (Table I).

The V_{th} -variation with $\sigma_{V_{th}} \propto t_{OX} \cdot N^{1/4} \cdot (1/W L_{eff})^{1/2}$ according to [6], [7] is calculated with the technology parameters given in Table I and scaling the impurity concentration N and the channel width W with $1/s$ [9] and s , respectively. For gates comprising small transistors and operated at $V_{DD} =$

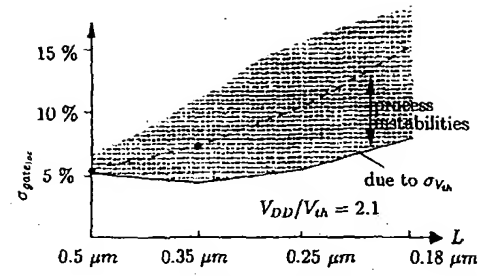


Fig. 8. Expected increase of the local delay variations $\sigma_{gate_{loc}}^r$ (dashed line for future technologies partitioned into effects according to doping variation (solid line) and process instabilities. A slight decrease for the doping variation of a 0.35 μm technology is observed. This is due to the usage of a dual work function technology with the p-channel devices as surface channel device. This results in reduced V_{th} variations. The dots show experimental results.

$2.1 \times V_{th}$ the trend of local delay variations is shown in Fig. by the solid line. For a 0.18 μm technology, $\sigma_{gate_{loc}}^r$ increases solely due to the physical effect of doping variations by a factor of two (5 \rightarrow 10%).

The dots show experimental results. The gap between predicted variations and measurements for the 0.35 μm technology can be explained by additional effects due to process instabilities. Especially for the more advanced technologies uncorrelated variations due to instabilities like e.g. transistor width deviations because of proximity effects [6] have to be taken into account.

The estimated contribution of this effect is shown symbolically by the shaded area. For our process, its estimated trend given by the dashed line. The resulting delay variation amounts to $\sigma_{gate_{loc}}^r = 15\%$ for a 0.18 μm technology.

The authors would like to point out that the solid line presents the lower limit of variations which exist due to unavoidable doping variations.

V. IMPACT OF LOCAL DELAY VARIATIONS ON PERFORMANCE AND YIELD

Local delay variations have to be considered in the design of low voltage circuits, where the impact gets significant compared to the effect of standard global variations. According to the results of the previous sections, circuits having short critical paths will suffer most from the influence of local delay variations. Short critical paths exist, e.g., in highly pipelined circuits.

Taking into account the global parameter variations on the gate delay values of the critical path follow a normal probability distribution function (pdf) $N_{gate_{glo}}(\bar{d}_{gate}, \sigma_{gate_{glo}}^r)$ where usually a 3 $\sigma_{gate_{glo}}^r$ -value of 50% is used. With the distribution function $F_{path_{crit}}$ of the critical path, the design can determine, e.g., the maximum applicable clock frequency $1/t_{clk}$ to achieve a specified yield.

In the case of low voltage circuits, an additional safety margin to account for the local effects has to be introduced. Uncorrelated gate delay variations result in a broader distribution $N_{gate_{glo+loc}}(\bar{d}_{gate}, \sigma_{gate_{glo+loc}}^r)$, with $\sigma_{gate_{glo+loc}}^r = \sqrt{(\sigma_{gate_{glo}}^r)^2 + (\sigma_{gate_{loc}}^r)^2}$. With $N_{gate_{glo+loc}}$ the path delay

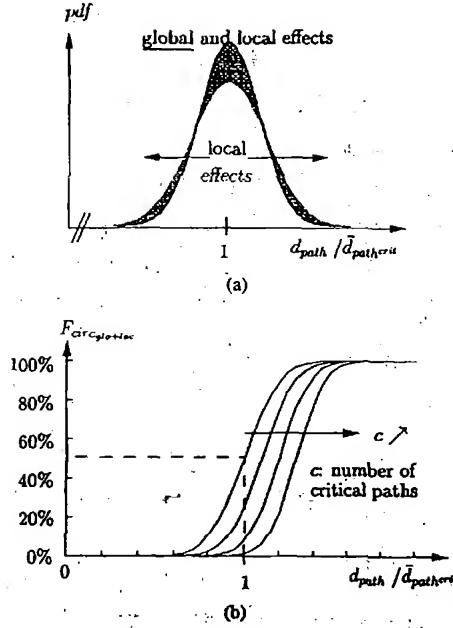


Fig. 9. (a) Additional uncorrelated gate delay variations lead to a broadened pdf of the yield. (b) Due to local effects the yield depends also on the number of critical paths which results in a shift of the distribution function $F_{\text{circ}_glo+loc}$.

pdf $N_{\text{path}_glo+loc}(\bar{d}_{\text{path}}, \sigma_{\text{path}_glo+loc}^r)$ comprising n gates can be calculated according to (26) in the Appendix with $\sigma_{\text{path}_glo+loc}^r = \sqrt{(\sigma_{\text{gate}_glo}^r)^2 + [(\sigma_{\text{gate}_loc}^r)^2/n]}$ (see Fig. 9).

In the case of one critical path the corresponding distribution function $F_{\text{path}_crit}^r$ governs the yield of the circuit.

If the local effects are small ($\sigma_{\text{gate}_loc}^r \approx 0\%$), this distribution represents also the yield for a circuit having c critical paths. In that case, the yield is independent of c . With $\sigma_{\text{gate}_loc}^r > 0\%$, the c critical paths show uncorrelated delay variations and yield is reduced [Fig. 9(b)] compared to $\sigma_{\text{gate}_loc}^r \approx 0\%$.

Instead of the pdf for a single path

$$f(d_{\text{path}}) = \frac{1}{\sqrt{2\pi} \cdot \sigma_{\text{path}_glo+loc}} \cdot e^{-0.5[(d_{\text{path}} - \bar{d}_{\text{path}})^2 / (\sigma_{\text{path}_glo+loc})^2]} \quad (8)$$

the pdf of a set of c critical paths has to be used given as the joint pdf of the path delays with

$$f(\bar{d}_{\text{path}}) = \frac{1}{[(2\pi)^c \cdot |C|]^{1/2}} \cdot e^{-(1/2) \cdot (\bar{d}_{\text{path}} - \bar{d}_{\text{path}})^T C^{-1} (\bar{d}_{\text{path}} - \bar{d}_{\text{path}})} \quad (9)$$

where $C_{(c \times c)}$ is the covariance matrix. To approximate the corresponding distribution function $F_{\text{circ}_glo+loc}(t_{\text{clk}}) = \int_{-\infty}^{t_{\text{clk}}} f(\bar{d}_{\text{path}})$ Monte Carlo simulations are used.

Examples of circuits with a high number of critical paths are data path structures, especially multiplier units. Depending on the architecture for a 32 bit multiplier based on carry save

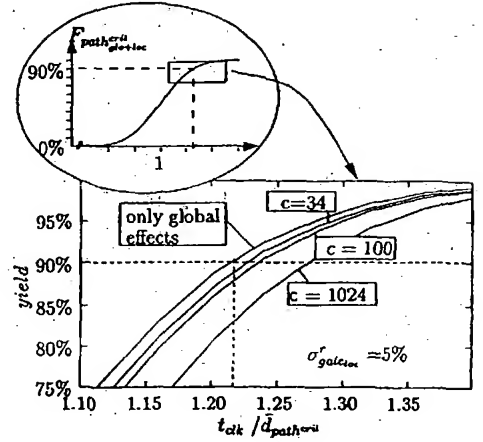


Fig. 10. Parametric yield normalized to the nominal path delay for sets of $c = 1, 34, 100$, and 1024 critical paths comprising a logic depth of $n = 132, 68, 36$, and 8 , respectively, for regarding global and local effects ($\sigma_{\text{gate}_loc}^r \approx 5\%$) and for regarding only the global effects. The inset shows the entire distribution function of $N_{\text{path}_glo+loc}$.

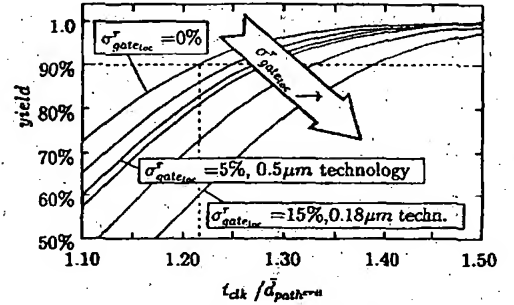


Fig. 11. Small section of the distribution functions for the path delays normalized to the nominal delays are shown for an example with $c = 1024$ paths for values of $\sigma_{\text{gate}_loc}^r$ between 0 and 15%.

adders the number c of critical paths can reach values of up to $32 \times 32 = 1024$. For that case, a logical depth of $n = 8$ gates is obtained for each critical path. This represents the situation for one full adder cell within each pipeline stage, comprising four gates for the full adder block, two gates for the AND-gate, and two gates for each latch.

For $\sigma_{\text{gate}_loc}^r = 5\%$ and using unit delays the distribution functions of the path delays normalized to the nominal delay $\bar{d}_{\text{path}_crit}$ are shown in Fig. 10 for values of $n = 8$ and $c = 1024$. For comparison the functions for $c = 1, 34$, and 100 are included and the parametric yield regarding only global effects is also given. For a scenario of 1024 critical paths with the logic depth of $n = 8$, the clock period has to be set to $1.27 \times \bar{d}_{\text{path}_crit}$ to reach a parametric yield of 90% instead of $1.21 \times \bar{d}_{\text{path}_crit}$ regarding only the global effects. For values of $\sigma_{\text{gate}_loc}^r = 15\%$ ($0.18 \mu\text{m}$ technology at $V_{DD} = 2.1 \times V_{th}$ in Fig. 11) the clock period increases significantly to $1.39 \times \bar{d}_{\text{path}_crit}$. Even for $\sigma_{\text{gate}_loc}^r = 3\%$ ($0.5 \mu\text{m}$ at $V_{DD} = 2.9 \times V_{th}$) the clock period amounts to $1.25 \times \bar{d}_{\text{path}_crit}$ (Fig. 11). The values for a 90% parametric yield are shown in Table II.

This results demonstrate that the impact of local delay

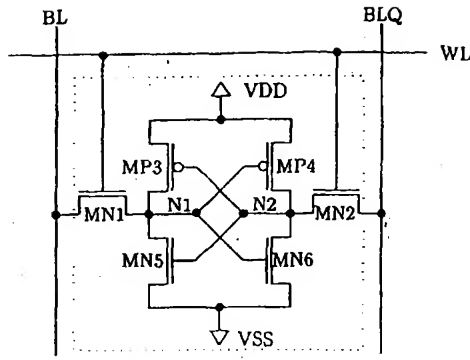


Fig. 12. Standard six transistor SRAM cell. For reading, the precharged bitlines BL or BLQ have to be discharged via MN1 or MN2. For MN1 and MN2 a width of two times the minimal gate length is used.

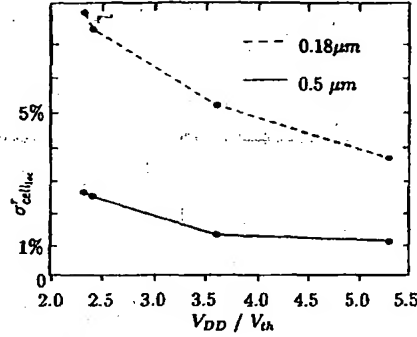


Fig. 13. Delay variations $\sigma_{cell,loc}^r$ of SRAM cell delay d_{cell} due to local parameter variations. Results for a $0.5 \mu m$ (simulated) and for a $0.18 \mu m$ (estimated) technology are given.

variation on yield evaluations depends strongly on the path spectrum, i.e., the number of critical paths and their logic depth.

A further example for functional blocks with a regular structure are embedded SRAM's. Bit failures due to V_{th} -variations can occur for four transistor cell structures and are investigated in [10]. For low power applications, six transistor cell [11] structures are used (Fig. 12) which are immune to soft errors and where the static power dissipation is low. Another characteristic of a SRAM is the access time for reading data. This is determined by the critical path comprising the address decoder, the cell, and the sense amplifier. Due to the small device dimensions, the most sensitive part to local parameter variations is the cell delay d_{cell} . The delay d_{cell} denotes the time for discharging the precharged bitlines BL or BLQ via transistor MN1 and MN5 or MN2 and MN6 respectively (Fig. 12). SPICE simulations show, that d_{cell} and its variations are clearly dominated by the properties of the transfer transistors MN1 and MN2. The influence of local parameter variations on the distribution function of d_{cell} is investigated for $3\sigma_{gate,lo}^r = 50\%$, $V_{DD} = 2.1V_{th}$, and W_{MN1} equals two times the minimum gate length L . Using the experimental determined V_{th} -variations of the NMOS transistor with the specific gate area, the delay variations $\sigma_{cell,loc}^r$ are determined by means of SPICE simulations (see

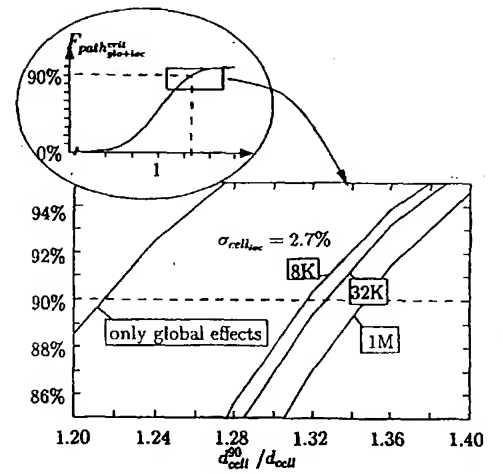


Fig. 14. Distribution functions for d_{cell}^{90} normalized to d_{cell} of SRAM's with 8K, 32K, and 1M are shown with local effects ($\sigma_{cell,loc}^r = 2.7\%$ corresponding to a $0.5 \mu m$ technology) and with regarding global effects only.

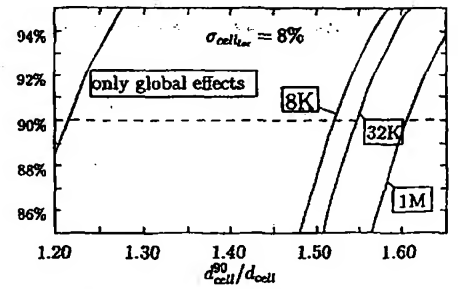


Fig. 15. Distribution functions for d_{cell}^{90} normalized to d_{cell} of SRAM's with 8K, 32K, and 1M are shown with local effects $\sigma_{cell,loc}^r = 8\%$, corresponding to a $0.18 \mu m$ technology) and with regarding global effects only.

Fig. 13). The cell delay d_{cell}^{90} , needed for yield of 90%, is calculated for different sizes of SRAM's, where the number of critical paths equals two times the size of the memory. For $0.5 \mu m$ technology with $\sigma_{cell,loc}^r = 2.7\%$, the ratio of d_{cell}^{90}/d_{cell} for sizes of 8K, 32K, and 1M is shown in Fig. 14. In the case of a 1M SRAM, d_{cell}^{90} amounts to $1.35 \times d_{cell}$ compare to $1.21 \times d_{cell}$ when local effects are neglected. For future technologies the effects increase significantly. For a $0.18 \mu m$ technology, d_{cell}^{90} is calculated to $1.60 \times d_{cell}$ to achieve the 90% yield for a 1M SRAM (Fig. 15). The exact values are given in Table III.

As the results shown above demonstrate an increased influence of local delay variations with decreasing logic depth an increasing number of critical paths, local effects have to be taken into account in designing pipelined circuits.

Neglecting register delay, with doubling the number of pipeline stages the critical path length is halved and the corresponding number of critical paths is doubled. Both lead to an increased effect of local delay variations on the performance.

As an example, the performance gain is determined for the 32-bit multiplier assuming 1, 2, 4, 8, and 16 pipeline stages with the critical paths comprising 68, 36, 20, 12, and

TABLE II
RATIO OF CLOCK PERIOD TO CRITICAL PATH DELAY (AND THE INCREASED CLOCK PERIOD IN PERCENTAGE
COMPARED TO REGARDING ONLY THE GLOBAL EFFECTS) TO ACHIEVE A PARAMETRIC YIELD OF 90%

$\sigma_{gate,loc}^r$	$c = 34, n = 68$		$t_{clk}/\bar{d}_{path,cris}$ for c critical paths: $c = 100, n = 36$		$c = 1024, n = 8$	
0%	1.21	(0%)	1.21	(0%)	1.21	(0%)
3% ($0.5 \mu m @ 2.9 V_{th}$)	1.21	(0%)	1.22	(0.8%)	1.25	(3.3%)
5% ($0.5 \mu m @ 2.1 V_{th}$)	1.23	(1.7%)	1.24	(2.5%)	1.27	(5.0%)
15% ($0.18 \mu m @ 2.1 V_{th}$)	1.25	(3.3%)	1.27	(5.0%)	1.39	(14.9%)

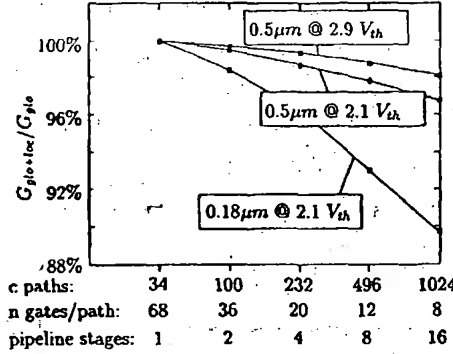


Fig. 16. Percentage of performance gain achieved for pipelining regarding local delay variations $G_{glo+loc}$ compared to regarding only global effects G_{glo} . For values of $\sigma_{gate,loc}^r = 3, 5, 15\%$ the example of pipelining 32 paths with $n = 68$ gates into $c = 1024$ paths comprising eight gates is shown.

TABLE III
RATIO OF THE CELL DELAY d_{cell}^{90} TO d_{cell} (AND THE INCREASE
IN PERCENTAGE COMPARED TO REGARDING ONLY THE GLOBAL
EFFECTS) TO ACHIEVE A PARAMETRIC YIELD OF 90%

$\sigma_{cell,loc}^r$	$c = 2 * 8 K$	$c = 2 * 32 K$	$c = 2 * 1M$
0%	1.21 (0%)	1.21 (0%)	1.21 (0%)
2.7% ($0.5 \mu m @ 2.1 V_{th}$)	1.32 (9%)	1.33 (10%)	1.35 (12%)
8.0% ($0.18 \mu m @ 2.1 V_{th}$)	1.52 (26%)	1.55 (28%)	1.60 (32%)

gates. The performance gain $G_{glo+loc}$ with $\sigma_{gate,loc}^r = 3\%, 5\%$, and 15% normalized to the performance gain G_{glo} regarding only global effects is shown in Fig. 16. For a $\sigma_{gate,loc}^r = 5\%$ the performance gain is reduced by 4%, for $\sigma_{gate,loc}^r = 15\%$ the reduction amounts to 10%. Therefore, the effect of local variations has to be checked when the optimum number of pipeline stages for a high performance architecture is determined.

VI. CONCLUSION

In this work we have presented the influence of local delay variations due to V_{th} -variations on path delays and on the design of low voltage digital circuits.

Motivated by SPICE simulations and analytical results which show relative gate delay variations due to local V_{th} -variations of up to $\sigma_{gate,loc}^r = 5\%$ of inverters with minimal feature size operating at low supply voltage, a test circuit for the determination of path delay variations was developed.

Measurements, taken for paths with different logic depths, verified the increase of the relative path delay variations for reduced logic depths, smaller dimensions, and reduced supply voltages. Circuits fabricated with a reduced threshold voltage

show reduced delay variations when operating at the same V_{DD} .

Due to the local variations, yield and performance of low voltage circuits depend on the number and the logic depth of the critical paths. Therefore, the influence on the yield and the performance was investigated for pipelined architectures and SRAM's. A scenario of 1024 critical paths with a logic depth of $n = 8$ gates shows already significant effects for an actual $0.5 \mu m$ technology. For example for $\sigma_{gate,loc}^r = 5\%$, the ratio of clock period to the nominal critical path delay is increased from 1.21 to 1.28 to achieve a parametric yield of 90% compared to designs operating at conventional supply voltages, where the yield is determined by highly correlated global gate delay variations. For SRAM circuits, local parameter variations have a significant impact on the cell delay. A scenario for three different SRAM's was investigated, showing that even for a current $0.5 \mu m$ technology in a 1 M SRAM the specification of the critical path delay has to be increased by 12% of the cell delay to achieve a yield of 90%.

In scenarios for future technologies the influence of V_{th} -variations grows. For a $0.18 \mu m$ technology, the margin to achieve the 90% yield will rise from 1.21 to 1.39.

Generally, for all signals, which are critical to skew like clock drivers or input stages of latches or flip flops, no minimal devices should be used. For the dynamic testing of circuits, the test of a single critical path is not sufficient any more to guarantee the function of the total circuit or to classify the circuit to a certain performance window. The tradeoff for a reduced V_{th} between higher performance and increased power dissipation due to leakage is extended to the aspect of reduced gate delay variations.

Further research will be done for low voltage memories which can comprise a high number of critical paths and for techniques with dynamic threshold voltage, to evaluate the influence of the V_{th} -variations.

APPENDIX

A. Variances of Gate and Path Delays

The probability distribution function pdf of the gate delay can be modeled by a normal distribution function $N(m_1, \sigma^2)$ fully characterized by its mean value $m_1 = \bar{d}_{gate}$ and its variance $\sigma^2 = (\sigma_{gate})^2$ by

$$f(d_{gate}) = \frac{1}{\sqrt{2\pi} \cdot \sigma_{gate}} \cdot e^{-0.5[(d_{gate} - \bar{d}_{gate})^2 / (\sigma_{gate})^2]} \quad (10)$$

The pdf of a *path* comprising n gates corresponds to the linear combination of the n pdfs of the gate delays. The average path delay is given by

$$\bar{d}_{\text{path}} = \sum_i^n d_{\text{gate}_i}. \quad (11)$$

With the symmetrical covariance matrix C

$$C = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ c_{21} & \dots & c_{2n} \\ \dots & \dots & \dots \\ c_{n1} & \dots & c_{nn} \end{bmatrix} \quad \text{and} \quad (12)$$

$$c_{ij} = \sigma_{\text{gate}_i} \cdot \sigma_{\text{gate}_j} \cdot \rho_{ij}$$

$$c_{ii} = \sigma_{\text{gate}_i}^2$$

ρ_{ij} : correlation between two gates

the variance of the path can be expressed as

$$(\sigma_{\text{path}})^2 = \sum_i^n \sum_j^n \sigma_{\text{gate}_i} \cdot \rho_{ij} \cdot \sigma_{\text{gate}_j} \quad (13)$$

$$= \sum_i^n \sum_j^n c_{ij} = \underline{e}^T \cdot C \cdot \underline{e} \quad [12]. \quad (14)$$

For $\rho_{ij} = \rho$, i.e., the intergate correlation is the same for all gates, and with the relative gate delay variation

$$\sigma_{\text{gate}}^r = \sigma_{\text{gate}} / \bar{d}_{\text{gate}}. \quad (15)$$

Equation (13) can be simplified to

$$\begin{aligned} (\sigma_{\text{path}}^2) &= \sum_i^n \sum_j^n \sigma_{\text{gate}_i} \cdot \rho \cdot \sigma_{\text{gate}_j} \\ &+ \sum_i^n (1 - \rho) (\sigma_{\text{gate}_i})^2 \\ &= \sum_i^n \sum_j^n \sigma_{\text{gate}_i}^r \cdot \bar{d}_{\text{gate}_i} \cdot \rho \cdot \bar{d}_{\text{gate}_j} \cdot \sigma_{\text{gate}_j}^r \\ &+ (1 - \rho) \sum_i^n (\sigma_{\text{gate}_i}^r)^2 (\bar{d}_{\text{gate}_i})^2 \\ \sigma_{\text{path}} &= \sigma_{\text{gate}}^r \sqrt{\rho (\bar{d}_{\text{path}})^2 + (1 - \rho) \sum_i^n (\bar{d}_{\text{gate}_i})^2} \\ &\text{for } \sigma_{\text{gate}_i}^r = \sigma_{\text{gate}}^r. \end{aligned} \quad (16)$$

Of interest are the two cases $\rho = 0$ for *local*, i.e., uncorrelated variations ($\sigma_{\text{gate}}^r = \sigma_{\text{gate}_{\text{loc}}}^r$) and $\rho = 1$ for the occurrence of only *global*, i.e., correlated variations ($\sigma_{\text{gate}}^r = \sigma_{\text{gate}_{\text{glo}}}^r$).

$$\rho = 0: \quad \sigma_{\text{path}} = \sigma_{\text{gate}_{\text{loc}}}^r \sqrt{\sum_i^n (\bar{d}_{\text{gate}_i})^2} \quad (17)$$

$$\sigma_{\text{path}}^r = \sigma_{\text{gate}_{\text{loc}}}^r / \sqrt{n} \text{ for } \bar{d}_{\text{gate}_i} = \bar{d}_{\text{gate}} \quad (18)$$

$$\rho = 1: \quad \sigma_{\text{path}} = \sigma_{\text{gate}_{\text{glo}}}^r \bar{d}_{\text{path}} \quad (19)$$

$$\sigma_{\text{path}}^r = \sigma_{\text{gate}_{\text{glo}}}^r \quad (20)$$

The first case clarify, that with $\rho = 0$ the relative path delay variation is reduced for longer paths (for identical gates with

$1/\sqrt{n}$). For the latter case, the relative path delay variation are constant as they are usually considered in the design with conventional supply voltages.

With both global and local variations, the pdf of the *gate delay* is given by the linear combination of $N_{\text{glo}}(\bar{d}_{\text{gate}}, \sigma_{\text{gate}_{\text{glo}}})$ and $N_{\text{loc}}(0, \sigma_{\text{gate}_{\text{loc}}})$ to $N_{\text{glo+loc}}(\bar{d}_{\text{gate}}, \sigma_{\text{gate}_{\text{glo+loc}}})$ with $\sigma_{\text{gate}_{\text{glo+loc}}}$ according to (14) and (15)

$$\begin{aligned} (\sigma_{\text{gate}_{\text{glo+loc}}})^2 &= (\sigma_{\text{gate}_{\text{glo}}})^2 + (\sigma_{\text{gate}_{\text{loc}}})^2 \\ \sigma_{\text{gate}_{\text{glo+loc}}} &= \bar{d}_{\text{gate}} \sqrt{(\sigma_{\text{gate}_{\text{glo}}}^r)^2 + (\sigma_{\text{gate}_{\text{loc}}}^r)^2}. \end{aligned} \quad (21)$$

The calculation of the variance of the *path delay* regarding local as well as global effects is demonstrated for an example with two gates. For that case, the correlation matrix $C_{m \times m}$ can be expressed with $m = 2n$ as

$$C_{\text{path}} = \begin{bmatrix} c_{11} & \dots & c_{12} \\ \dots & \dots & \dots \\ c_{21} & \dots & c_{22} \end{bmatrix} \quad (22)$$

where the submatrixes (simplified for gate 1 with $\sigma_{1\text{glo}}^2$ and $\sigma_{1\text{loc}}^2$ and for gate 2 with $\sigma_{2\text{glo}}^2$ and $\sigma_{2\text{loc}}^2$) are

$$C_{11} = \begin{bmatrix} \sigma_{1\text{glo}}^2 & 0 \\ 0 & \sigma_{1\text{loc}}^2 \end{bmatrix}, \quad C_{12} = C_{21} = \begin{bmatrix} \sigma_{1\text{glo}} \cdot \sigma_{2\text{glo}} & 0 \\ 0 & 0 \end{bmatrix},$$

$$C_{22} = \begin{bmatrix} \sigma_{2\text{glo}}^2 & 0 \\ 0 & \sigma_{2\text{loc}}^2 \end{bmatrix}$$

and the correlation coefficients between the gates are

$$\rho_{\text{glo glo}} = 1, \quad \rho_{\text{loc glo}} = 0, \quad \rho_{\text{glo loc}} = 0, \quad \text{and} \quad \rho_{\text{loc loc}} = 0. \quad (23)$$

The variance $(\sigma_{\text{path}_{\text{glo+loc}}})^2$ of the path can be calculated according to (14) with

$$(\sigma_{\text{path}_{\text{glo+loc}}})^2 = \sum_i^m \sum_j^m c_{ij} C_{\text{path}}, \quad m = 2. \quad (24)$$

For $\sigma_{1\text{glo}} = \sigma_{2\text{glo}} = \sigma_{\text{glo}}$ and with (17) and (19) this expression is simplified to

$$\sigma_{\text{path}_{\text{glo+loc}}} = \sqrt{(\sigma_{\text{glo}}^r \bar{d}_{\text{path}})^2 + \sum_i^{n=2} (\bar{d}_{\text{gate}_i} \cdot \sigma_{\text{gate}_{\text{loc}_i}}^r)^2}. \quad (25)$$

For a path with n gates and for $\bar{d}_{\text{gate}_i} = \bar{d}_{\text{gate}}$ and $\sigma_{\text{gate}_{\text{loc}_i}}^r = \sigma_{\text{gate}_{\text{loc}}}^r$ the standard deviation of a path can be simplified to

$$\begin{aligned} \sigma_{\text{path}_{\text{glo+loc}}} &= \bar{d}_{\text{path}} \sqrt{(\sigma_{\text{gate}_{\text{glo}}}^r)^2 + (\sigma_{\text{gate}_{\text{loc}}}^r)^2 / n} \\ &= \bar{d}_{\text{path}} \cdot \sigma_{\text{path}_{\text{glo+loc}}}^r. \end{aligned} \quad (26)$$

For example, with increasing logic depth, the impact of the local variations is reduced.

ACKNOWLEDGMENT

The authors would like to thank E. Ruderer for his cooperation.

DOCKET NO: LL-10217

SERIAL NO: _____

APPLICANT: J. Berthold et al.

LERNER AND GREENBERG P.A.

P.O. BOX 2480

HOLLYWOOD, FLORIDA 33022

TEL. (954) 925-1100

A. KLASIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
 IPK 7 G06F17/50

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
 IPK 7 G06F

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

INSPEC, IBM-TDB, PAJ, EPO-Internal

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
Y	<p>EISELE M ET AL: "The impact of intra-die device parameter variations on path delays and on the design for yield of low voltage digital circuits"</p> <p>1996 INTERNATIONAL SYMPOSIUM ON LOW POWER ELECTRONICS AND DESIGN, MONTEREY, CA, USA, 12-14 AUG. 1996,</p> <p>Bd. 5, Nr. 4, Seiten 360-368, XP002243764</p> <p>IEEE Trans. Very Large Scale Integr. (VLSI) Syst. (USA), IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Dec. 1997, IEEE, USA</p> <p>ISSN: 1063-8210</p> <p>Absatz '0005!</p> <p>appendix</p> <p>Abbildungen 9-16</p> <p>Tabellen 2,3</p> <p style="text-align: center;">---</p> <p style="text-align: center;">-/--</p>	1-6



Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen



Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen

A Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

E älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

L Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

O Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

P Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

T Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

X Veröffentlichung von besonderer Bedeutung, die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

Y Veröffentlichung von besonderer Bedeutung, die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

Z Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

11. Juni 2003

Absendedatum des internationalen Recherchenberichts

24/06/2003

Name und Postanschrift der internationalen Recherchenbehörde

Europäisches Patentamt, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx 31 651 epo nl,
 Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Guingale, A

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
Y	<p>BOWMAN K A ET AL: "Impact of extrinsic and intrinsic parameter fluctuations on CMOS circuit performance" IEEE JOURNAL OF SOLID-STATE CIRCUITS, AUG. 2000, IEEE, USA, Bd. 35, Nr. 8, Seiten 1186-1193, XP002243765 ISSN: 0018-9200 das ganze Dokument</p> <p>-----</p> <div style="border: 1px solid black; padding: 5px; margin-top: 20px;"><p>Docket # <u>L&L-10217</u> Applic. # _____ Applicant: <u>J. Berthold et al</u> Lerner and Greenberg, P.A. Post Office Box 2480 Hollywood, FL 33022-2480 Tel: (954) 925-1100 Fax: (954) 925-1101</p></div>	1-6